Evaluating ChatGPT's Performance in the EU*US eHealth Work Foundational Curriculum Using the HITCOMP Self-Assessment Quiz

Chun-Wei LIU[a,b] , Zu-Chun LIN[a], and Malcolm KOO[a1]

ªDept. of Nursing, Tzu Chi University of Science and Technology, Hualien, Taiwan
ᵇMedical Intensive Care Unit, Mennonite Christian Hospital, Hualien, Taiwan
ORCiD ID: Chun-Wei Liu https://orcid.org/0009-0007-7477-8551
Zu-Chun Lin https://orcid.org/0000-0001-7994-0617
Malcolm Koo https://orcid.org/0000-0002-9242-9167

Abstract. This study investigated the performance of OpenAI's Chat Generative Pre-trained Transformer (ChatGPT) in responding to the EU*US eHealth Work Foundational Curriculum. This curriculum, a collaborative effort between European and U.S. institutions, provides an extensive framework for eHealth learning. The assessment involved 321 questions from the online Health Information Technology Competencies (HITCOMP) self-assessment quiz. Using GPT-3.5 model, the study presented each question three times to assess ChatGPT's consistency. Findings revealed an accuracy of 70.7%, indicating a reasonable grasp of eHealth topics, although performance was uneven across the 21 modules. These results underscore ChatGPT's potential in health information technology education and highlight the need for further model enhancements to fully encompass eHealth competencies.

Keywords. Artificial intelligence, ChatGPT, eHealth competency, health information technology

1. Introduction

The EU*US eHealth Work Foundational Curriculum, a key advancement in health information technology (HIT) education, was collaboratively developed by European and U.S. institutions. It offers a comprehensive framework for eHealth skills and knowledge, encompassing 10 clusters, 21 modules, and 60 units in its online Health Information Technology Competencies (HITCOMP) resource (http://hitcomp.org/).
The Chat Generative Pre-trained Transformer (ChatGPT), developed by OpenAI, is an advanced language model capable of responding to complex queries. Its potential in managing the vast amount of information prevalent in the healthcare sector makes it a significant asset. Although tested in various professional healthcare examinations, its efficacy in eHealth competencies has not been examined previously [1–3]. This study aimed to assess the eHealth competency of ChatGPT based on the questions available in the HITCOMP self-assessment quiz.

---

[1] Corresponding Author: Malcolm Koo, E-mail: mkoo@ems.tcust.edu.tw

2. Methods

The study utilized 321 questions from the HITCOMP self-assessment quiz. ChatGPT's performance was assessed by presenting each question in triplicate to evaluate its response consistency. ChatGPT's responses were compared to the correct answers. The study was conducted on Dec 1, 2023, using the GPT-3.5 model via ChatGPT website.

3. Results

ChatGPT correctly answered 227 of the 321 HITCOMP questions (70.7%) across three attempts, indicating a reasonable understanding of the eHealth topics covered in the curriculum. However, its performance was not uniformly across all modules, with less consistency in areas, such as Data Protection and Security in eHealth (module 17) (42.9% correct) (Figure 1).
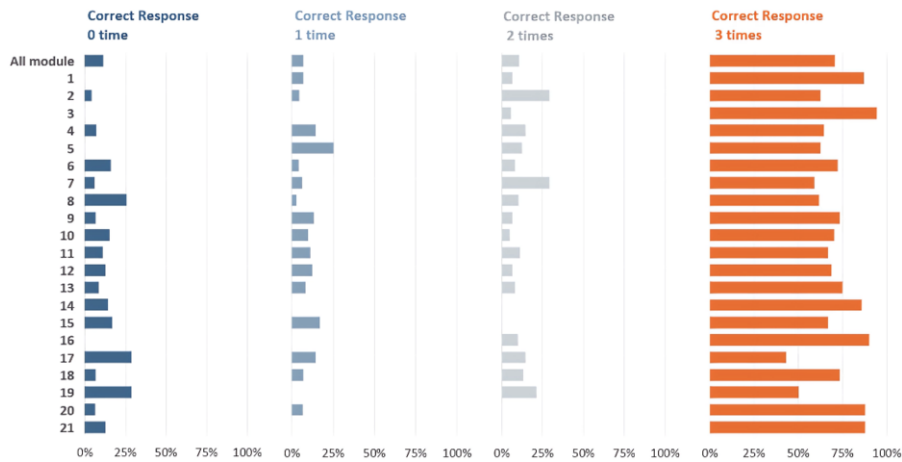

Figure 1. Results of ChatGPT responses to 321 questions of the HITCOMP self-assessment quiz. Complete titles are available at http://hitcomp.org/education/.

4. Conclusions

The findings from this study highlight the potential of artificial intelligence tools, such as ChatGPT, in enhancing learning of health information technology. They also identify areas where further development of the AI model is needed to ensure comprehensive coverage and understanding of all eHealth competencies.

References

[1] Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ. 2023;9:e45312. doi:10.2196/45312

[2] Bhayana R, Krishna S, Bleakney RR. Performance of ChatGPT on a radiology board-style examination: Insights into current strengths and limitations. Radiology. 2023;307(5):e230582. doi:10.1148/radiol.230582

[3] Huang H. Performance of ChatGPT on registered nurse license exam in Taiwan: A descriptive study. Healthcare (Basel). 2023;11(21):2855. doi:10.3390/healthcare11212855